JANUARY 2014, NO 2

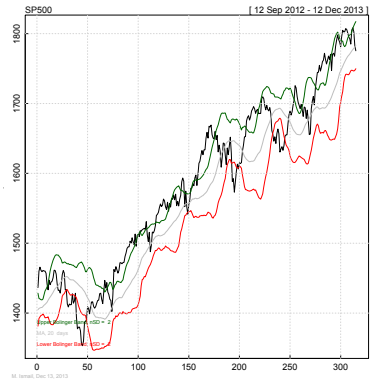
Research & Insights

Analytical Research Ltd is an independent research company that provides professional quantitative and qualitative research for both the public and private sectors.

We are experienced in mathematical trading and financial modelling such as econometric analysis, structured derivatives pricing, portfolio construction and risk management.

Analytical Research Ltd is registered in England and Wales, No. 6933835.
Station House, Connaught Road, Surrey, GU24 0ER.

Trend Analysis of Equity Data



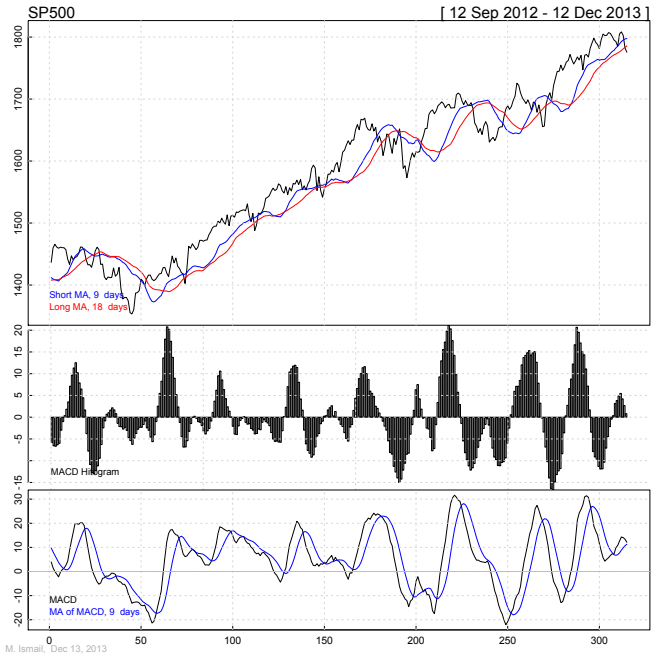
When analysing equity data for the purpose of trading, usually two approaches come to mind; time series models such as AR, ARCH, GARCH and trend analysis. The first approach is usually used for in-sample analysis, for example: estimating the mean and volatility, or for investigating causality between different time series, e.g. SP500 and FTSE100. Trend analysis, on the other hand, is focused on short-term prediction of equity data movement for the goal of achieving profit. Even though, in theory, predicting the market contradicts with the 'Efficient Market Hypothesis' which states it is impossible to beat the market, this approach is probably the most used by market practitioners, including large investment banks and hedge funds for the purpose of algorithmic trading. There is a vast amount of academic literature that explains why trend analysis should still work; the least to mention is that, not all traders are rational. In this release, we will mention briefly technical analysis, then we shall move on to explain one of the modelling approaches used by Analytical Research Ltd, to utilise different information for the purpose of black box trading.

Technical Analysis

It is the generic term that describes pattern recognition tools used by traders and market practitioners to create a short-term forecast. Many tools come under this umbrella ranging from simple charts and statistical indicators, up to discriminant analysis, neural networks and genetic algorithms. Trend followers may make use of high frequency tick data, to trade within a very short time frame, others aim to predict the market for longer period of time, i.e. several days into the future. It is always the case that several indicators are used together to reach a trading decision.

Statistical Indicators

They are broadly classified into moving average indicators, directional movement indicators, and oscillators. However, it is hard to draw lines between these classifications. For example, for some people MACD, or Moving Average Convergence-Divergence is a directional indicator, that helps to stay long in up-trends and short in down-trends, see (page 2, left figure). For others it is referred to as an oscillator. Another example is ADX, Average Directional Index, which measures trend strength regardless to its direction, see (page 2, right figure). Another example is Bollinger bands: described as a standard deviation channel, as it marks a moving average ± 2 standard deviation, see the top figure. This is useful for measuring change in volatility, as the outer lines depart when volatility is high. For a practitioner, charts reflect what has happened, Indicators reveal the balance of power between bulls and bears, and oscillators help catch turning points, within a trading range.



Statistical Model

To design an algorithm for black box trading, we use several indicators and combine their collective output via statistical modelling. The strategy we explain is based on fitting a probit model to a response variable with an ordinal outcome (Short/Out/Long).

Ideal trading positions are modelled in relation to a set of indicators X. Although the model could be estimated using an MLE-based method, we prefer to do the estimation within a Bayesian framework to obtain a distribution for every estimated coefficient, even for quantities not directly estimated in the model. This offers advantages in relation to better residuals' diagnostics, and easier test of hypotheses for confounding effects, and facilitates further analysis. Using the data augmentation approach [1], we assume a latent continues variable Z_i distributed as $N(X_i^T\beta, 1)$, and we observe the trading position Y_i where,

$$y_i = \begin{cases} short & if \quad -\infty = \gamma_0 < z_i <= \gamma_1 \\ out & if \quad \gamma_1 < z_i <= \gamma_2 \\ long & if \quad \gamma_2 < z_i < \gamma_J = \infty \end{cases}$$

The coefficient vector β and the bin boundaries γ are unknown. The full posterior density for β and γ could then be written as follows:

$$\pi(\beta, \gamma|Y) \propto \pi(\beta, \gamma) \prod_{i=1}^{N} \sum_{j=1}^{J} 1(y_i = j) \times$$
$$\left[\phi\left(\gamma_j - X_i^T\beta\right) - \phi\left(\gamma_{j-1} - X_i^T\beta\right)\right]$$

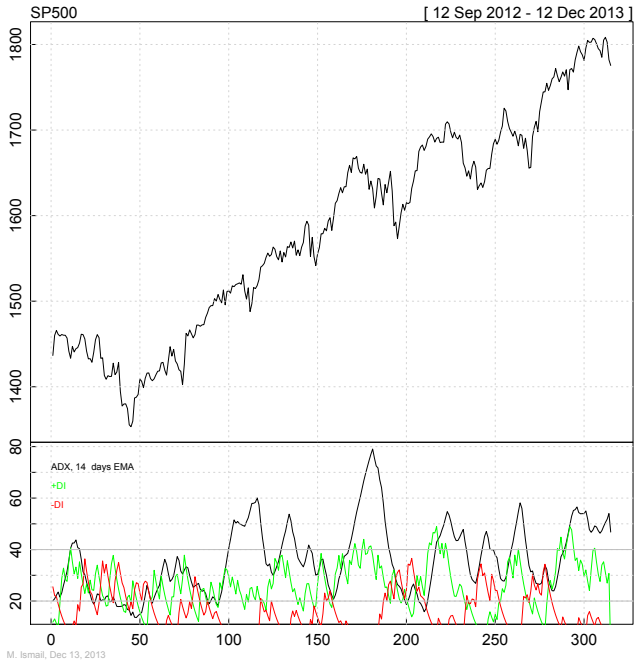
where $\pi(\beta, \gamma)$ is the prior distribution. See [1] for how to make this model fully Bayesian by choosing a suitable prior, and defining the full conditional densities for β,γ,Z_i which are necessary for implementing a Gibbs sampler.

Other modelling approaches can be used such as neural networks, or genetic algorithms. Using these approaches, the historical data is divided into training, testing and prediction sets. Attention is always given to data mining. It is claimed that a one layer network is a good approximation to any nonlinear function, however, neural networks are highly parameterised, and appear to predict poorly outside the sample boundaries.

Testing

We use large number of stocks for design, training, and testing. For each stock a historical data set that covers several years is used. Allocating the first 40% of the data set for algorithms' design, it allows the next 60% for testing, prediction, and bench marking.

Many indicators are evaluated prior to deciding the trading rules and modelling. We also invest considerable efforts in the design of visualisation techniques and tools for the purpose of monitoring and optimising the trading algorithms. We use and test these visualisation tools in our early design stage and they prove to be efficient for highlighting pitfalls and missing opportunities.



We hope you enjoyed reading this release and hope to see you soon!

References

[1] Albert, J. and Chib, S (1993), "Bayesian Analysis of Binary and Polychotomous Response Data," *Journal of the American Statistical Association*, 88, 669-679.